METHODS OF STREAMFLOW DATA ANALYSIS

DEFINITIONS

Discharge (Q): The *volume of water per unit time* that passes a specified point on a stream. Discharge is conventionally measured in cubic feet per second (ft^3/sec or cfs) or cubic meters per second (m^3/sec or cms). $Q = w\ d\ v$ where w = water width d = mean water depth v = mean water velocity

Rating curve (or **stage-discharge curve):** A curve relating *stage* (water height) to water discharge at a point on the stream channel.

Water year: For hydrologic purposes, we conventionally use the *water year*, which runs from October of one year through September of the next, so that we do not break up the winter storm flows between years. (For example, water year 1990 extends from Oct. 1 1989 through Sept. 30 1990.)

Annual flood: The *annual flood* on a stream is the highest instantaneous peak discharge of the water year.

Flood magnitude: The size of a flood peak in discharge units (e.g., ft^3/sec or m^3/sec).

Flood recurrence interval (or **return period):** The average time in years between flood events equal to or greater than a specified magnitude.

Flood-frequency curve: A graph showing the relationship between flood magnitude and their recurrence interval for a specified site.

Mean daily discharge: The average discharge of any specified calendar day (midnight to midnight). It is calculated by taking the total *volume* of water discharged during that day and dividing by 86400, the number of second in a day

Flow-duration curve: A graph showing the percentage of time specified mean daily discharges are equalled or exceeded.

Mean annual discharge (Q_{av}): The discharge that would have to flow constantly to equal the *volume* of water discharged by that stream over the entire period of years record. Q_{av} is the total volume of water discharged by the stream during the period of record divided by the number of seconds in that period. It is more easily calculated by averaging all the individual mean daily discharges.

Hydraulic geometry: The hydraulic geometry of a stream is the set of relations which show how width, depth, velocity, and cross-sectional area of flow vary with increasing discharge. These relations take the form of power functions (straight lines on logarithmic graph paper).

General Procedures for Flood Frequency Calculation from Gaging Data

INTRODUCTION

A *flood-frequency curve* at a point on a stream shows how often flood discharges of different sizes (magnitudes) will be equalled or exceeded. In combination with a rating curve and a topographic map, flood-frequency curves can be used to predict how often various areas are likely to be inundated.

The *recurrence interval* (T_r) of a flood is a *statistical* measure of how often a flood of a given magnitude is likely to be equalled or exceeded. Specifically, the "fifty-year flood" is one which will, *on the average*, be equalled or exceeded once in *any* fifty-year period. It does *not* mean that it occurs every fifty years.

The *probability* that a flood of specific size will be equalled or exceeded in any given year is given by:

$$P = \frac{1}{T_r}$$

The fifty-year flood has once chance in fifty of occurring in any specified year, that is, its probability is 1/50 (P = 0.02 or 2%).

Flood discharges of varying recurrence intervals are symbolized as $\mathbf{Q_{Tr}}$, where $\mathbf{T_r}$ is the appropriate recurrence interval. For example, the fifty-year recurrence interval flood (or simply the fifty-year flood) is written $\mathbf{Q_{50}}$.

There are three chief approaches to constructing a flood frequency curve at a gaged site:

a. Empirical (non-parametric) flood-frequency curve -- essentially an eyeball fit of the relation between estimated recurrence interval and flood magnitude.

b. Parametric flood-frequency curve -- a curve created by fitting the data to a specified probability distribution, such as the lognormal, extreme value, or log-Pearson Type III distribution

c. Flood-frequency curve derived from regional relations, e.g., from a regression on drainage area, mean annual discharge, mean annual rainfall, relief, mean annual flood, channel width, etc.

PROCEDURE FOR DEVELOPING AN EMPIRICAL FLOOD-FREQUENCY CURVE

1. Compile a list of *annual floods.*

   Set up a flood frequency computation table as follows:

Water Year	Date	Peak Q ft^3/s	Rank, M	T_r yr

   From published gaging data (e.g., USGS or DWR data publications) under "momentary maximum" or "annual maximum" find the largest flood (maximum instantaneous peak discharge) of each *water* year. Enter that value in ft^3/sec in the "Peak Q" column.

2. Rank the discharges

   When all the floods have been entered in the computation table--one for each year--rank the discharges in order from largest to smallest. Let the largest flood have rank M=1. The smallest flood will have rank M=N, where N is the number of years for which we have flood data. In the case of a tie, give the tied events different , but adjacent, ranks.

3. Compute the recurrence interval

   Compute the *recurrence interval* (or *return period*) of each flood using the formula:

   $$T_r = \frac{(N+1)}{M}$$

   M = rank

   N = total number of floods

   The units of T_r are years.

4. Plot the discharges on flood-frequency paper

   Plot each flood discharge versus its T_r on either arithmetic or logarithmic extreme-value flood frequency paper. Fit a *smooth curve* through the points. Note that this is a best-fit-by-eye curve, *not* follow the dots. You are trying to extract a general trend from the data. Because we do not have much data to define the uppermost end of the curve, we should not try slavishly to fit the line through the uppermost points--instead, we need to look at the overall trend of the curve (c.f. Dunne & Leopold p. 305-313.) Note: In fitting the curve, don't give too much weight to the position of the largest 2 or 3 floods if the sample size is small (n < 25). They are probably plotting to the left of where they would correctly plot if we had a longer record.

   This graph is the flood-frequency relation for the gaging station, based on the N years of data available to us. The relation would change if we had more years of data available.

5. It is often a good idea to draw confidence bands around the flood-frequency curve. A procedure for doing this is described on p. 308-09 of Dunne & Leopold.

6. From this curve you can read off the estimated flood discharge corresponding to different recurrence intervals (e.g. 1.5-yr, 5-yr, 10-yr, 50-yr floods). To get the floods for larger recurrence intervals (ones beyond the limits of your plotted data), you will have to cautiously extrapolate the flood-frequency curve. There will be substantial error in these estimates when the period of record is short.

PROCEDURE FOR DEVELOPING A PARAMETRIC FLOOD-FREQUENCY CURVE

1. Compile a list of *annual floods.* (See step 1 in previous section for details.)
2. Take the common logs ($\log_{10}$) of the flood discharges.
3. Compute the mean $\bar{\mathbf{x}}$ and standard deviation $\mathbf{s}$ of the original data. Then compute the mean $\overline{\mathbf{x_l}}$, standard deviation $\mathbf{s_l}$, and coefficient of skewness $\mathbf{g_l}$ of the log-transformed data.
4. The general form of the equation defining the floods of different recurrence intervals is

   $$\mathbf{Q_{Tr}} = \bar{\mathbf{x}} + \mathbf{K_T}\,\mathbf{s}$$

   where K_T is called the frequency factor and depends on the probability distribution assumed for the floods

5. **Lognormal distribution**

For a lognormal fit, the defining equation is: $\mathbf{Q_{Tr} = 10^{\overline{x_l} + K_{TL}\, s_l}}$

The values of K_T to use in the equation are listed in the table below:

recurrence interval T_r yr	frequency factor K_{TL}
1.5	-0.439
2	0.000
5	0.842
10	1.282
25	1.751
50	2.054
100	2.326
200	2.576
500	2.878

The computed discharges can be plotted versus recurrence interval on semi-log or flood frequency paper and fitted with a smooth line.

6. **Type I extreme-value (Gumbel) distribution**

For an extreme-value fit, the defining equation is: $\mathbf{Q_{Tr} = \overline{x} + K_{TG}\, s}$

where $\mathbf{K_{TG} = -\frac{\sqrt{6}}{\pi}\{0.577 + \ln[\ln T_r - \ln(T_r - 1)]\}}$

Values of the frequency factor are given in the table below

recurrence interval T_r yr	frequency factor K_{TG}
1.5	-0.523
2	-0.164
5	0.720
10	1.305
25	2.044
50	2.592
100	3.137
200	3.679
500	4.395

The computed discharges can be plotted versus recurrence interval on flood frequency paper and fitted with a smooth line.

Computation and Plotting of Flow-Duration at a Gaging Site

INTRODUCTION

The *flow-duration curve* (or *duration curve*) of a stream is based on daily *mean* discharges (not peak flows) and shows the percentage of time that a given daily mean discharge is equalled or exceeded. The duration curve is extremely useful in water-supply studies, especially of low flows. For example, it can show us the percent of time that the discharge on a stream is below some critical value, such as the amount that needs to be diverted to a powerhouse, or for public water supply, or to dilute effluent. In a geological sense, the flow-duration curve can be used in conjunction with a sediment-transport curve or dissolved-solids rating curve to compute the amount of sediment or dissolved solids leaving a drainage basin.

Discharges corresponding to a particular *exceedence frequency* (or *duration*) are symbolized as $\mathbf{q_d}$ where $\mathbf{d}$ is the percent of time that the specified mean daily discharge is equalled or exceeded. For example, $\mathbf{q_{35}}$ is the mean daily discharge that is equalled or exceeded 35% of the time; 65% of the time the flows are less than this value.

PROCEDURE

1. Determine the size classes for tabulating the daily discharges

   Examine the range of flow values present in the data. Determine the number of log cycles (powers of 10) the data spans, and then for tallying data use the class intervals suggested in the table below.

Range in daily discharge				
1 log cycle	2 log cycles	3 log cycles	4 log cycles	5 log cycles
10	10	10	10	10
11	12	15	15	15
12	14	20	20	20
13	17	25	30	30
14	20	30	40	50
15	25	40	50	70
16	30	50	70	100
18	35	60	100	150
20	40	80	150	etc.
22	45	100	etc.	
24	50	150		
26	60	etc.		
28	70			
30	80			
33	100			
36	120			
40	etc.			
45				
50				
55				
60				
65				
70				
75				
80				
90				
100				

NOTE: Table shows sequence of numbers for five ranges in discharge. Locate decimal point and starting discharge to suit conditions. In general, use cycle closest to observed range. Where considerable data have been tabulated at a station by using other rules for class intervals, additional data would normally be tabulated by using the class intervals previously established. Source: USGS Water-Supply Paper 1542-A

2. Construct a table of size classes (e.g., 100-150 cfs, 70-100 cfs, etc.) for tallying daily discharges; arrange the categories from largest to smallest. An example is shown below.

Flow-Duration Compilation Table

Q	number of days with Q< upper class limit but ≥ lower class limit													% time	Cum% time ≥ lower limit
	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	June	July	Aug	Sept	Total Days		
100 –150															
70 –100															
50 –70															
30 –50															
20 –30															
15 –20															
10 –15															
7 –10															

3. Tally the discharges into the size categories

   From appropriate sources of mean daily discharge data (e.g., USGS or DWR compilations), tally (by months) the number of flows in each discharge size category into the boxes on the flow-duration compilation table. The categories work as follows: Q's < 150 cfs but ≥ 100 cfs (i.e. 100 ≤ Q < 150) would be tallied into boxes in the row labeled 100 –150; Q's < 100 cfs but ≥ 70 cfs would be tallied into the row labeled 70 –100, etc.

   After you have finished tallying all the discharges, add them across to get the total number of days in each size category. Enter these totals in the "Total Days" column.

4. Determine the percent of time and cumulative % of time

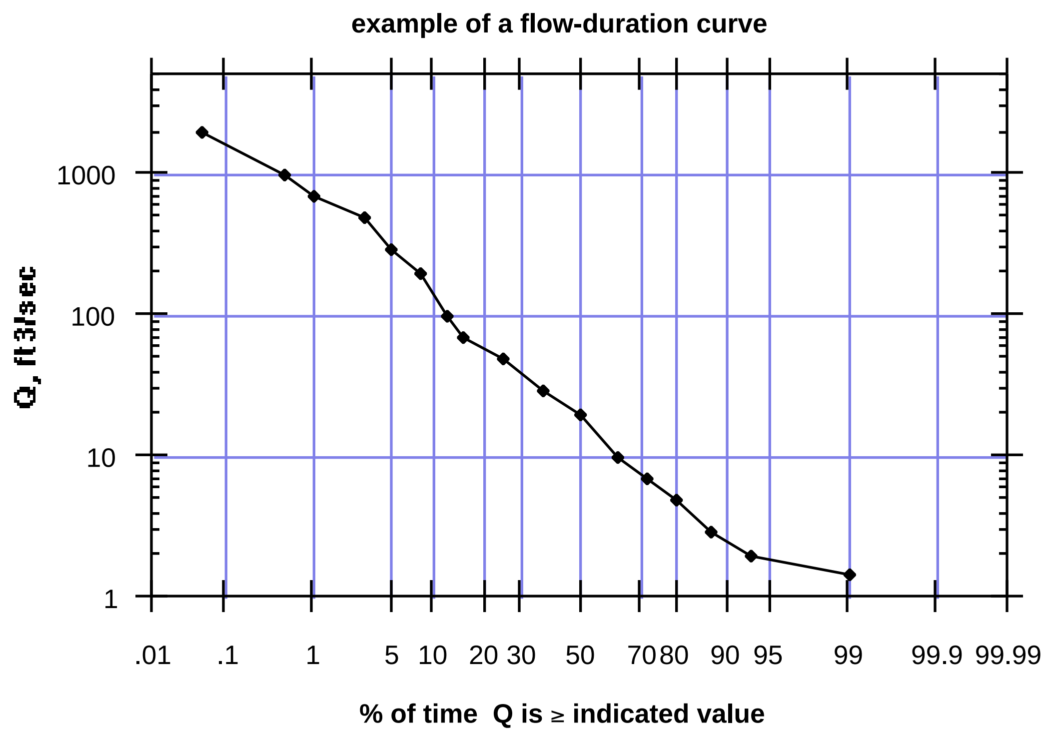
   Determine the *percent of time* each flow category represents by dividing the total days in each class by the total number of days encompassed by our data. For example if we had 2 years of record (730 days) and 37 days in the 100-150 cfs category, then (37/730) x 100% or approximately 5.1% of the time would be occupied by flows between 100 and 150 cfs.

   Add the individual percents sequentially downward (from larger to smaller discharges) to find the *cumulative percent of time* that discharges *equal or exceed the lower limit* of each size category. The very lowest category will have a value of 100% since all the discharges will be greater than it.

5. Plot the curve on log-probability paper

   On log-probability paper plot the water discharge (in ft^3/s or m^3/s) corresponding to the *lower* boundary of each class versus the "cumulative % of time" that discharges equal or exceed that category. The discharges are plotted on the log scale (vertical axis), and the cumulative % on the probability scale (horizontal axis). It is customary to have 0.01% at the left-hand edge of the graph and 99.99% at the right-hand edge. Note that it is impossible to actually plot 0% or 100% on the paper, but you can safely plot your lowest flow at 99.99%.

6. Draw a smooth follow-the-dots curve *through all the points*. This is the flow-duration curve.



Regional Generalization of Flow-Duration Curves

INTRODUCTION

We are often faced with a situation where we need a flow-duration curve for an ungaged stream or for an ungaged site on a stream. Described below are two techniques for estimating duration curves for such sites.

PROCEDURES

1. We can regionalize a curve by dividing the curve ordinates (discharge) by the *drainage area* (A_d) of the stream. This is particularly useful when we are dealing with smaller streams and wish to estimate the duration curve at one point on the stream given a curve developed elsewhere on the stream. This technique relies on the basic *assumption* that discharge is directly proportional to drainage area, i.e., that all parts of the basin contribute the same amount of discharge per unit area. If this assumption does not hold, this technique will give inappropriate results.

2. An alternative regionalization is to divide the curve ordinates by the *mean annual discharge* (Q_{av})of the stream at the gaging site. This yields a dimensionless flow-duration curve. This technique is more general than dividing only by drainage area, as it integrates the effects of drainage area, rainfall, vegetation, and other basin characteristics which affect mean annual flow. I have found this approach particularly useful for transferring curves from one basin to another, especially if the basins differ appreciably in rainfall. If we can estimate the mean annual discharge on a stream (which we usually can from rainfall–drainage area relations, or from a water budget), then we can synthesize a duration curve for the stream from the regionalized one.

3. In both cases, to develop the regionalized curves we compute the duration curves for a number of gaged streams in the area of interest, and divide their ordinates as described in 1 or 2 above. We plot the regionalized data for all the streams on log-probability paper, and draw a best-fit curve through the points. This is our *regional duration curve*. You will generally find that if you are dividing by Q_{av} the points from all the curves will lie close together for the larger (i.e., less frequent) flows, and will spread apart at the low-flow end. This reflects the different low-flow characteristics of each basin, which are strongly controlled by local geology, topography, and vegetation.

Determining At-A-Station Hydraulic Geometry

INTRODUCTION

The *hydraulic geometry* of a stream is the set of relations that show how width, depth, and velocity vary with discharge either at a point on the stream or in a downstream direction. The hydraulic geometry at a given cross-section is called the *at-a-station* hydraulic geometry. The hydraulic geometry shows us how the stream adjusts its characteristics in order to carry the water discharge and sediment loads supplied to it. To determine the hydraulic geometry we need several sets of discharge measurements together with the water width, mean depth, and mean velocity at the time of the measurement.

PROCEDURE

1. On logarithmic graph paper plot:

width vs. discharge	(w vs Q)
mean depth vs. discharge	(d vs. Q)
mean velocity vs. discharge	(v vs. Q)

   Plot Q on the horizontal axis, and w, d, and v on the vertical axis.

2. Draw a separate best-fit straight line through each group of points. Do this in pencil because you will probably have to adjust these lines.

3. A straight line on logarithmic paper is a power function. The straight lines you have drawn define relations of the form:

   $w = aQ^b$ $\qquad$ $d = cQ^f$ $\qquad$ $v = kQ^m$

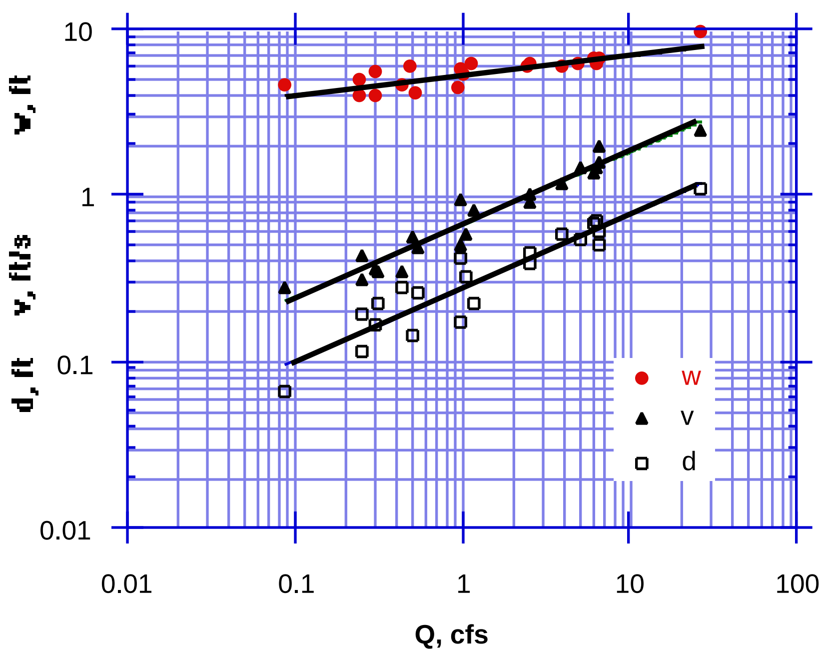
   where b, f, and m are the slopes of the lines, and a, c, and k are the values of w, d, and v respectively when Q = 1.0.

   Since by definition $Q = wdv$, then $Q = (aQ^b)(cQ^f)(kQ^m) = ack\, Q^{b+f+m}$

   This can only hold if $ack = 1$ *and* $b+f+m = 1$.

   That is, the slopes of the lines must sum up to 1, and the product of the intercepts (at Q = 1) must equal 1.

   Thus you will have to fiddle with your best-fit lines, adjusting their slope and intercept, until these conditions are approximately met. You want b+f+m to deviate from 1.0 by no more than 0.02 (i.e., $0.98 \le b+f+m \le 1.02$) and ack should be within 0.05 of 1 (i.e., $0.95 \le ack \le 1.05$). Of course, the closer you can get them to 1, the better.



$w = 5.34\, Q^{0.118}$
$d = 0.280\, Q^{0.443}$
$v = 0.672\, Q^{0.439}$